

**PERANCANGAN SISTEM LLM AGENT BERBASIS MODEL
PERILAKU DALAM PENGEMBANGAN SOCIAL MEDIA BOT
INSTAGRAM UNTUK MENGIMITASI PERILAKU
INFLUENCER**

**Laporan Proposal Tugas Akhir - Capstone
Pengembangan *Social Media Bot* Instagram Terkait Pariwisata
Berbasis LLM dan Data Mining**

**Disusun sebagai syarat kelulusan mata kuliah
IF4091/Penulisan Proposal**

**Oleh
JUAN CHRISTOPHER SANTOSO
NIM : 13521116**



**PROGRAM STUDI TEKNIK INFORMATIKA
SEKOLAH TEKNIK ELEKTRO & INFORMATIKA
INSTITUT TEKNOLOGI BANDUNG
JANUARI - 2025**

**PERANCANGAN SISTEM LLM AGENT BERBASIS MODEL
PERILAKU DALAM PENGEMBANGAN SOCIAL MEDIA
BOT INSTAGRAM UNTUK MENGIMITASI PERILAKU
INFLUENCER**

Laporan Proposal Tugas Akhir - Capstone
Pengembangan *Social Media Bot* Instagram Terkait Pariwisata
Berbasis LLM dan Data Mining

Oleh
JUAN CHRISTOPHER SANTOSO
NIM : 13521116

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung

Bandung, 6 Januari 2025

Mengetahui,

Pembimbing I,

Pembimbing II,

Dr.techn. Saiful Akbar, S.T., M.T.

NIP. 19740509 199803 1 002

Dr. Agung Dewandaru, S.T., M.Sc.

NIP. 122110013

LEMBAR IDENTITAS

TUGAS AKHIR CAPSTONE

Judul Proyek TA : Pengembangan *Social Media Bot* Instagram Terkait
Pariwisata Berbasis LLM dan Data Mining

Anggota Tim dan Pembagian Peran:

| No. | NIM | Nama | Peran |
|-----|----------|---------------------------|---------------------------------------|
| 1 | 13521100 | Alexander Jason | Perancangan Modul Perangkat Lunak |
| 2 | 13521116 | Juan Christopher Santoso | Perancangan Modul LLM |
| 3 | 13521139 | Nathania Calista Djunaedi | Perancangan Modul Penambangan Data |

Bandung, 6 Januari 2025

Mengetahui,

Pembimbing I,

Pembimbing II,

Dr.techn. Saiful Akbar, S.T., M.T.

NIP. 19740509 199803 1 002

Dr. Agung Dewandaru, S.T., M.Sc.

NIP. 122110013

DAFTAR ISI

| | |
|--|------------|
| DAFTAR ISI..... | i |
| DAFTAR LAMPIRAN | iii |
| DAFTAR GAMBAR..... | iv |
| DAFTAR TABEL | v |
| DAFTAR SINGKATAN..... | vi |
| DAFTAR ISTILAH | vii |
| BAB I PENDAHULUAN..... | 1 |
| I.1 Latar Belakang..... | 1 |
| I.2 Rumusan Masalah..... | 3 |
| I.3 Tujuan dan Ukuran Keberhasilan Pencapaian | 4 |
| I.4 Batasan Masalah | 4 |
| I.5 Metodologi..... | 5 |
| BAB II STUDI LITERATUR | 7 |
| II.1 Media Sosial | 7 |
| II.2 Instagram | 8 |
| II.3 <i>Social Media Bot</i> | 8 |
| II.4 Pariwisata di Indonesia | 9 |
| II.5 <i>Deep Learning</i> | 10 |
| II.6 LLM Agents | 11 |
| II.7 <i>Large Language Models</i> (LLMs) | 13 |
| II.7.1 <i>Knowledge Base</i> (KB)..... | 14 |
| II.7.2 <i>Language Models</i> | 14 |

| | | |
|---|--|-----------|
| II.7.3 | <i>Pre-trained Models</i> | 15 |
| II.7.4 | Teknik <i>Tuning</i> | 16 |
| II.7.5 | Teknik <i>Prompting</i> | 19 |
| II.7.6 | <i>State Memory Flows</i> | 20 |
| II.7.7 | <i>Self-Detection</i> | 20 |
| II.8 | Model Perilaku Pengguna | 23 |
| BAB III ANALISIS MASALAH DAN RANCANGAN SOLUSI PENGEMBANGAN LLM PADA <i>SOCIAL MEDIA BOT</i> INSTAGRAM..... | | 24 |
| III.1 | Deskripsi Umum Persoalan <i>Capstone</i> | 24 |
| III.2 | Analisis Masalah | 24 |
| III.3 | Analisis Solusi | 25 |
| III.4 | Deskripsi Solusi..... | 26 |
| BAB IV RENCANA PELAKSANAAN..... | | 27 |
| IV.1 | Jadwal | 27 |
| IV.2 | Risiko..... | 29 |
| DAFTAR PUSTAKA | | 31 |

DAFTAR LAMPIRAN

| | |
|---|----|
| Lampiran A. Rencana Umum Proyek | 35 |
| Lampiran B. Spesifikasi Kebutuhan Sistem..... | 38 |
| Lampiran C. Rancangan Sistem..... | 40 |

DAFTAR GAMBAR

| | |
|--|----|
| Gambar II.1. <i>Platform</i> media sosial dengan tingkat pengunjungan paling tinggi di Indonesia (We Are Social, 2024) | 8 |
| Gambar II.2 Visualisasi dari arsitektur <i>deep learning</i> (Azawii dkk., 2019) | 10 |
| Gambar II.3 Komponen sebuah LLM <i>agent</i> (Bansal, 2024) | 11 |
| Gambar II.4 Arsitektur dari proses LLMs (Omiye dkk., 2023) | 14 |
| Gambar II.5 Visualisasi capaian dari <i>language model pre-training</i> (Naveed dkk., 2024) | 16 |
| Gambar II.6 Visualisasi proses pemodelan LLMs (Capella, 2024) | 17 |
| Gambar II.7 Ilustrasi dari proses RAG (Parthasarathy dkk., 2024) | 18 |
| Gambar II.8 Visualisasi <i>consistency-based component</i> (Zhao dkk., 2024) | 21 |
| Gambar II.9 Visualisasi <i>verbalization-based component</i> (Zhao dkk., 2024) | 22 |
| Gambar III.1 Diagram arsitektur dari <i>social media bot</i> Nusava | 26 |
| Gambar IV.1 <i>Gantt chart</i> pengerjaan proposal tugas akhir | 27 |
| Gambar IV.2 <i>Gantt chart</i> pengerjaan tugas akhir | 28 |
| Gambar A.1 Diagram posisi <i>social media bot</i> | 35 |
| Gambar C.1 Arsitektur diagram <i>social media bot</i> | 40 |
| Gambar C.2 Alur kerja sistem LLM <i>agent</i> | 43 |

DAFTAR TABEL

| | |
|--|----|
| Tabel IV.1 Deskripsi <i>task</i> dalam pengerjaan proposal tugas akhir | 27 |
| Tabel IV.2 Deskripsi <i>task</i> dalam pengerjaan tugas akhir..... | 29 |
| Tabel IV.3 Daftar risiko dan rencana mitigasi dari risiko..... | 29 |
| Tabel A.1 Daftar pihak dan interaksinya dengan <i>bot</i> | 35 |
| Tabel B.1 Spesifikasi kebutuhan fungsional sistem..... | 38 |
| Tabel B.2 Spesifikasi kebutuhan non-fungsional sistem | 39 |
| Tabel C.1 Komponen utama <i>social media bot</i> | 40 |
| Tabel C.2 Komponen pendukung <i>social media bot</i> | 41 |
| Tabel C.3 Komponen sistem LLM <i>agent</i> | 43 |

DAFTAR SINGKATAN

| Singkatan | Kepanjangan Singkatan |
|------------------|--------------------------------|
| AI | <i>Artificial Intelligence</i> |
| KB | <i>Knowledge Base</i> |
| LLM | <i>Large Language Model</i> |
| LLMs | <i>Large Language Models</i> |
| LMs | <i>Language Models</i> |
| ML | <i>Machine Learning</i> |

DAFTAR ISTILAH

| Istilah | Makna Istilah |
|---|---|
| <i>Artificial intelligence</i> (arti: kecerdasan buatan) | Bidang keahlian yang mempelajari pengembangan komputer dan mesin sehingga mampu untuk melakukan pembelajaran dan berperilaku selayaknya kecerdasan manusia. |
| <i>Classifier</i> (arti: penggolong) | Sebuah algoritma yang bertujuan untuk melakukan klasifikasi atau pengelompokkan data ke dalam rentang atau kelas tertentu. |
| <i>Hallucinations</i> (arti: halusinasi) | Sebuah fenomena pada system LLMs dimana hasil konten yang dibangkitkan oleh sistem bersifat tidak relevan, dibuat-buat, dan tidak konsisten. |
| <i>Machine learning</i> (arti: pembelajaran mesin) | Bidang keahlian di dalam bidang <i>artificial intelligence</i> yang mempelajari mengenai pengembangan sebuah algoritma statistik menggunakan data |
| <i>Natural language</i> (arti: bahasa alami) | Bahasa yang secara alami digunakan oleh manusia dalam berinteraksi. Dalam hal ini, bahasa alami adalah bahasa yang memang secara fundamental dimengerti oleh manusia, contohnya Bahasa Indonesia, bukan bahasa untuk berinteraksi dengan sistem komputer seperti halnya bahasa pemrograman. |
| <i>Neural network</i> (arti: jaringan syaraf tiruan) | Metode pada bidang <i>artificial intelligence</i> dalam mengembangkan kecerdasan pada komputer untuk |

memproses data dengan basis pemrosesan pada otak manusia.

Prompt

(arti: perintah)

Sebuah perintah yang diberikan pengguna kepada sistem untuk melakukan suatu pekerjaan

BAB I

PENDAHULUAN

Implementasi LLMs dalam pengembangan *social media bot* adalah salah satu hal yang relatif baru untuk diimplementasikan. Meski penggunaannya yang masih belum umum, terdapat berbagai kebutuhan dari pengguna yang sebenarnya dapat dipenuhi dengan diimplementasikannya LLMs pada *social media bot*. Pada bab ini akan dipaparkan lebih lanjut terkait detail-detail pelaksanaan tugas akhir, mencakup alasan pelaksanaan, permasalahan yang diselesaikan, tujuan pelaksanaan, batasan dari pelaksanaan, dan metodologi yang dilakukan.

I.1 Latar Belakang

Industri pariwisata menjadi salah satu industri yang memiliki peran penting dalam pembangunan nasional (Rusyidi dan Fedryansah, 2018). Pada tahun 2016, pendapatan dari sektor pariwisata di Indonesia berkontribusi senilai Rp500,19 triliun dari *Product Domestic Bruto* (PDB) dan terus meningkat hingga saat ini. Dengan berbagai keindahan alam yang ada di Indonesia, tak sedikit wisatawan baik dalam negeri maupun mancanegara yang tertarik untuk menjelajahi tempat wisata yang ada di Indonesia.

Menurut data pada Badan Pusat Statistik (2024), jumlah wisatawan mancanegara yang datang ke Indonesia mencapai angka 11 juta turis dalam satu tahunnya. Data ini diperjelas dengan salah satu *sample* jumlah kunjungan wisatawan mancanegara ke Indonesia yakni 1,07 juta turis pada Bulan September tahun 2023 (Badan Pusat Statistik, 2023). Angka tersebut memang jauh lebih sedikit apabila dibandingkan dengan jumlah wisatawan lokal, yakni 192 juta turis dalam rentang 3 bulan. Namun, angka tersebut menunjukkan bahwa upaya terkait menaikkan jumlah wisatawan mancanegara masih terbuka lebar.

Dengan banyaknya destinasi wisata yang tersedia di Indonesia, pemilihan destinasi dan pengalaman wisata menjadi tantangan tersendiri bagi para turis yang akan mengunjungi Indonesia. Dalam hal ini, media sosial menjadi kakas yang berperan

penting bagi turis dalam menentukan rencana wisata. Menurut Chakraborty (2024), dalam salah satu media sosial, Instagram, terdapat lebih dari satu miliar *post* yang memiliki *hashtags* berkaitan dengan pariwisata. Tak hanya itu, 89% dari wisatawan menggunakan media sosial sebagai sumber inspirasi dalam melakukan kegiatan pariwisata, dengan media sosial yang digunakan diantaranya Instagram, Facebook, dan Pinterest. Terlebih lagi, media sosial dikatakan memiliki kontribusi hingga 41% dalam mempengaruhi wisatawan dalam melakukan pemesanan rencana wisata. Hal ini menjelaskan besarnya pengaruh media sosial dalam dunia pariwisata.

Seiring dengan tingginya tingkat ketergantungan turis dengan platform media sosial, media sosial menjadi salah satu sumber informasi yang digunakan turis untuk melakukan kegiatan pariwisata. Bahkan, tak jarang bagi turis untuk dipengaruhi dan terdorong untuk melakukan kegiatan wisata setelah melihat berbagai konten pariwisata di media sosial. Maka dari itu, selayaknya seorang *influencer*, dibutuhkan sebuah teknologi yang mampu membantu para turis dalam memenuhi kebutuhan mereka dan mendorong mereka dalam melakukan kegiatan pariwisata. Pada platform media sosial, fungsi ini dapat dilakukan oleh *social media bot*.

Social media bot adalah *bot* yang bertempat pada platform media sosial dan memiliki kapabilitas untuk berperilaku dan berinteraksi selayaknya seorang manusia (Grimme dkk., 2017). *Social media bot*, mengimitasi perilaku manusia, dapat melakukan berbagai aksi, beberapa diantaranya, *like*, *follow*, *post*, dan *comment*. Namun, *social media bot* seringkali dapat dengan mudah diidentifikasi dari cara *bot* tersebut menyampaikan pesan dan melakukan aksi yang terkesan kaku. Maka dari itu, *social media bot* masih terbuka dengan pengembangan yang lebih lanjut agar dapat menyerupai perilaku manusia dengan lebih baik.

Cara kerja dari *social media bot* umumnya dilandasi oleh sistem otomasi berbasis *script* dan algoritma (Oestreicher, 2024). Namun, untuk mengatasi permasalahan kesan kaku dari *bot*, *social media bot* perlu didasari oleh model yang mampu mengerti dan memahami bahasa yang digunakan oleh manusia. Model pengenalan bahasa ini lebih umum dikenal dengan *large language models* (LLMs).

LLMs yang digunakan dalam *social media bot* umumnya merupakan *pre-trained models* yang telah diberikan pembelajaran dasar untuk dapat memahami bahasa alami. Namun meski telah melewati pembelajaran dasar, sebuah LLM umumnya masih belum mumpuni apabila dibutuhkan untuk mengerjakan pekerjaan spesifik. Maka dari itu, untuk memenuhi kebutuhan ini, diperlukan proses kustomisasi untuk sistem LLM itu sendiri.

Kustomisasi sebuah LLM dapat dilakukan dengan berbagai cara. Menurut Belcic & Stryker (2024), dalam penerapan sebuah LLM, perlu dilakukan penyetelan (*tuning*) agar kualitas *pre-trained* model tetap terjaga. Di sisi lain, menurut Naveed dkk. (2024), performa sebuah model dapat ditingkatkan dengan melakukan penyesuaian pada instruksi (*prompt*) yang diberikan. Harapannya, proses kustomisasi dari sebuah LLM dapat meningkatkan performa dari LLM itu sendiri sehingga tercipta sebuah sistem LLM *agent*.

LLM *agents* adalah sistem berbasis AI yang memiliki kemampuan untuk menyelesaikan berbagai permasalahan kompleks yang membutuhkan urutan penalaran (Bansal, 2024). Dalam pelaksanaannya, LLM *agents* juga mampu untuk menyesuaikan respon mereka berdasarkan kondisi yang dibutuhkan. Tak hanya itu, LLM *agents* juga memiliki kapabilitas untuk menentukan pilihan secara otomatis dan melakukan pembelajaran secara mandiri. Sistem LLM *agents* inilah yang perlu diimplementasikan pada *social media bot* agar dapat mengimitasi perilaku manusia.

I.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, dapat diketahui bahwa *social media bot* digunakan untuk memenuhi peran *influencer*, yakni memenuhi kebutuhan pengguna untuk mendapatkan informasi secara efektif dan melakukan aksi yang sesuai dengan masukan pengguna. Dengan memenuhi persyaratan tersebut, harapannya, *bot* ini dapat sekaligus berfungsi untuk mendorong para pengguna untuk melakukan kegiatan pariwisata. Lantas, permasalahan ini menimbulkan sebuah pertanyaan terkait kebenaran pengimplementasian LLMs dapat meningkatkan performa imitasi *bot* terhadap perilaku manusia. Tak hanya itu,

perlu juga dikaji terkait proses kustomisasi LLMs yang dapat dilakukan sehingga dihasilkan model yang dapat menyerupai peran *influencer*.

Maka dari itu, rumusan masalah yang dibentuk berdasarkan latar belakang yang telah disampaikan adalah sebagai berikut:

1. Apakah *social media bot* dapat meningkatkan efektivitas dalam memenuhi kebutuhan pengguna terkait pariwisata?
2. Bagaimana cara merancang *social media bot* Instagram yang terotomasi dan dapat meniru perilaku seorang *influencer* di bidang pariwisata?
3. Apakah implementasi LLMs pada *social media bot* dapat meningkatkan imitasi *bot* terhadap perilaku manusia?
4. Bagaimana proses kustomisasi LLMs dapat dilakukan sehingga dihasilkan model yang menyerupai *influencer* pariwisata ?

I.3 Tujuan dan Ukuran Keberhasilan Pencapaian

Berdasarkan rumusan masalah yang telah dibentuk, tujuan dari tugas akhir ini adalah sebagai berikut:

1. Dihasilkannya sistem dalam *social media bot* yang mampu memenuhi kebutuhan *user* terkait pariwisata
2. Dihasilkannya sebuah *social media bot* yang terotomasi dan meniru perilaku seorang *influencer*
3. Peninjauan hasil dari implementasi sistem LLMs pada *social media bot* dalam mengimitasi perilaku manusia
4. Pengkajian terkait kustomisasi LLM dalam mengimitasi sosok *influencer* pariwisata di *social media*

I.4 Batasan Masalah

Batasan masalah dari tugas akhir ini dapat diuraikan sebagai berikut:

1. Data yang digunakan sebagai landasan bagi *knowledge base* sistem adalah data yang berkaitan dengan topik pariwisata di Provinsi Nusa Tenggara Barat dan Nusa Tenggara Timur.
2. Masukan yang diberikan kepada sistem berupa teks dan gambar. Di sisi lain, keluaran yang dihasilkan oleh sistem berupa teks. Pilihan aksi yang dapat dilakukan oleh *bot* adalah *like*, *follow*, *comment*, *reply*, dan *post*.
3. Bahasa yang digunakan dalam pengembangan, penggunaan, dan pembelajaran sistem, baik pengguna kepada sistem maupun sistem kepada pengguna, adalah Bahasa Inggris. Bahasa Inggris dipilih dengan pertimbangan fokus penggunaan *social media bot* yang diarahkan untuk digunakan oleh wisatawan mancanegara.

I.5 Metodologi

Metodologi pengerjaan tugas akhir ini dapat diuraikan sebagai berikut:

1. Tahap eksplorasi dan analisis

Tahap ini berisikan proses pencarian ide dan pendalaman dari topik yang dikerjakan untuk tugas akhir. Berikutnya, proses eksplorasi dilanjutkan dengan pelaksanaan analisis secara teoritis. Tahap ini juga yang menjadi landasan bagi pengerjaan pendahuluan pada Bab I dan landasan teoritis pada Bab II.

2. Tahap perancangan sistem

Tahap ini mencakup seluruh proses perencanaan dari pengembangan sistem *social media bot* baik dari penggunaan bahasa pemrograman, pemanfaatan *tech stack* dan model, serta pemilihan teknik. Tahapan ini didasari oleh landasan teoritis yang tertera pada Bab II.

3. Tahap implementasi sistem

Tahap ini berisikan seluruh proses penerapan sistem yang telah dirancang pada tahap perancangan sistem. Proses ini dimulai dari pembangunan sistem, baik dari awal maupun pengembangan sistem yang

sudah ada, hingga dihasilkan sebuah sistem yang dapat digunakan oleh pengguna.

4. Tahap evaluasi dan pengujian

Tahap ini mencakup proses pengukuran keberhasilan dari sistem yang telah diimplementasikan. Metode-metode pengujian, baik kualitatif dan kuantitatif, digunakan pada tahap ini untuk mengukur kualitas dan performa sistem.

5. Tahap penarikan kesimpulan

Tahap ini meliputi proses pengambilan pengetahuan (*knowledge*) dan wawasan (*insight*) yang disimpulkan dari hasil pengembangan sistem. Tahap ini merupakan tahap akhir sekaligus penutup dari pengerjaan tugas akhir yang menjawab tujuan dari pengerjaan tugas akhir.

BAB II

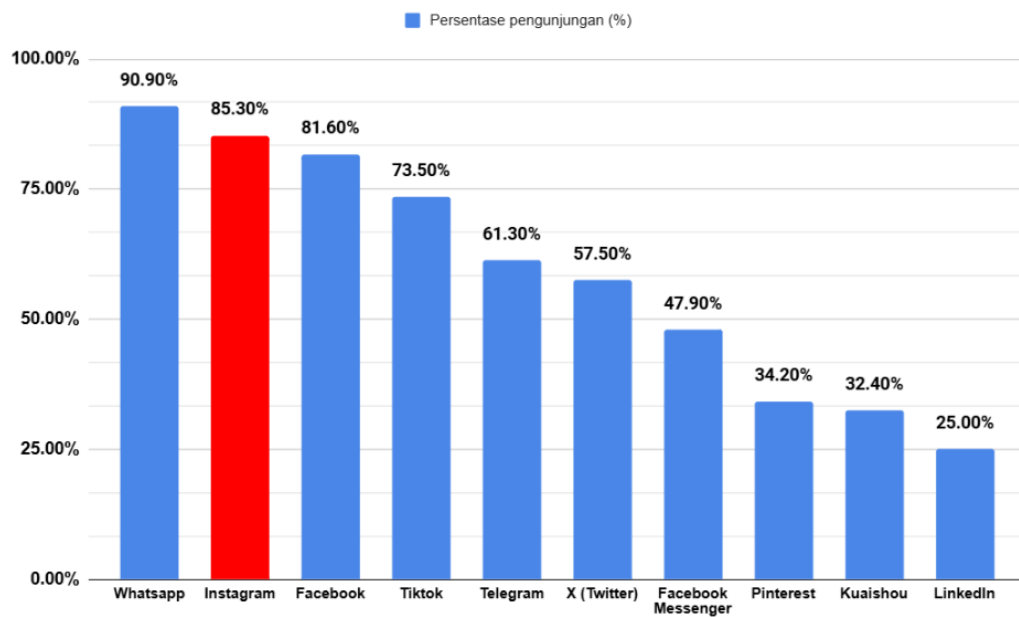
STUDI LITERATUR

Perancangan dari sistem LLM *agents* pada *social media bot* tidak dilakukan secara sembarangan. Studi literatur perlu dilakukan untuk memahami secara mendalam konsep-konsep teoritis dalam pelaksanaan tugas akhir ini. Tak hanya itu, eksplorasi terkait penelitian terdahulu perlu dilakukan guna melaksanakan tugas akhir ini secara optimal.

II.1 Media Sosial

Media sosial adalah media di internet yang memungkinkan pengguna merepresentasikan dirinya maupun berinteraksi dengan pengguna lain dan membentuk ikatan sosial secara *virtual* (Widada, 2018). Kegiatan sosial dalam hal ini mencakup: pengenalan (*cognition*), komunikasi (*communication*), dan kerjasama (*co-operation*). Menurut Widada, media sosial memberikan kontribusi babak baru dalam peradaban ilmu pengetahuan. Ilmu pengetahuan tidak lagi terpusat kepada golongan terpelajar tetapi tersebar kepada semua orang yang ikut terkoneksi dalam jejaring sosial.

Berdasarkan data yang dirilis oleh organisasi We Are Social (2024), per Januari 2024, terdapat 185,3 juta pengguna internet di seluruh Indonesia dengan 98,4% penggunaan tersebut digunakan untuk mengakses jejaring sosial (*social networks*). Dari 185,3 juta pengguna internet, terdapat 139 juta identitas yang terdaftar sebagai pengguna media sosial. Angka ini setara dengan 49.9% populasi yang ada di Indonesia pada tahun 2024. Platform media sosial dengan tingkat pengunjungan paling tinggi di Indonesia dapat ditampilkan sebagai berikut:



Gambar II.1. *Platform* media sosial dengan tingkat pengunjungan paling tinggi di Indonesia (We Are Social, 2024)

II.2 Instagram

Instagram adalah platform media sosial yang dikembangkan oleh Meta. Pada masa kini, Instagram menjadi salah satu platform media sosial yang memiliki pengguna paling banyak. Menurut Zote (2024), berdasarkan statistik Instagram pada tahun 2024, terdapat lebih dari 2 milyar pengguna aktif dari Instagram di seluruh dunia, dengan 60% dari pengguna tersebut berada pada rentang usia 18 hingga 34 tahun.

Selain memenuhi fungsinya sebagai media sosial, Instagram juga dapat digunakan sebagai media pelaksanaan usaha atau bisnis. Berdasarkan data Instagram pada tahun 2024, 70% dari pembeli potensial melakukan eksplorasi pada platform Instagram terlebih dahulu sebelum membeli sebuah produk atau layanan. Hal ini mencakup jasa dan layanan pariwisata yang informasinya tersebar luas pada platform Instagram.

II.3 Social Media Bot

Social bots adalah sebuah *user account* yang terotomasi baik sepenuhnya maupun sebagian sehingga mampu berpartisipasi dalam interaksi *online* (Magnus & Hakan,

2022). Di sisi lain, *social media bots* adalah *social bots* yang bertempat pada platform media sosial dan berperilaku selayaknya seorang manusia. Magnus & Hakan menambahkan bahwa *social media bots* berinteraksi dan memengaruhi manusia dengan tujuan politik, ideologi, maupun komersil. Pernyataan ini didukung oleh Grimme dkk. (2017) yang menyatakan bahwa *social bots* dianggap sebagai faktor yang berpengaruh sekaligus misterius dalam membuat opini. Menurut Grimme, tidak dapat dipungkiri bahwa *social bots* memiliki pengaruh besar dalam dunia sosial.

Berdasarkan Indusface (2024), *social media bots* dapat melakukan berbagai pekerjaan selayaknya seorang *user* di media sosial. Beberapa diantaranya adalah melakukan *posting* sebuah konten, melakukan *like*, memberikan *comment* pada sebuah *post*, melakukan *follow* pada akun lainnya, dan memberikan pesan pribadi atau *direct messages*. Bergantung kepada pemakaiannya, *social media bots* dapat digunakan untuk memudahkan berbagai pekerjaan. Namun bila disalahgunakan, *social media bots* juga dapat digunakan untuk melakukan aktivitas yang berbahaya, seperti penyebaran misinformasi, manipulasi opini publik, dan pelaksanaan *spam*.

II.4 Pariwisata di Indonesia

Industri pariwisata menjadi salah satu industri yang memiliki peran penting dalam pembangunan nasional (Rusyidi & Fedryansah, 2018). Pada tahun 2016, pendapatan dari sektor pariwisata di Indonesia menunjukkan perkembangan dan kontribusi yang terus meningkat yakni sebesar 4,03% dari *Product Domestic Bruto* (PDB) dengan nilai Rp500,19 triliun. Berkaca pada potensi tersebut, pengembangan pariwisata menjadi salah satu program unggulan dalam pembangunan nasional maupun daerah.

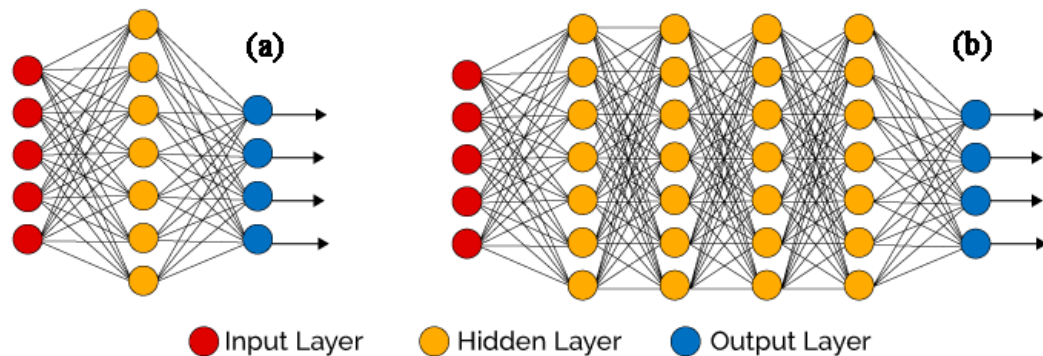
Menurut Setiawan (2015), salah satu bentuk pemberdayaan dari Indonesia, sebagai negara yang kaya akan sumber daya alam, adalah diciptakannya tempat sarana destinasi wisata. Hal ini menjadikan Indonesia sebagai negara dengan penuh potensi pada industri pariwisata. Untuk mengembangkan potensi yang ada,

dilakukan program pengembangan pariwisata berkelanjutan. Pariwisata berkelanjutan (*sustainable tourism*) dalam hal ini adalah pariwisata yang memperhitungkan dampak ekonomi, sosial budaya, dan lingkungan saat ini hingga masa depan (Wibowo & Belia, 2023). Dengan kata lain, sektor pariwisata adalah salah satu industri kunci bagi pengembangan Indonesia di masa yang akan datang.

II.5 Deep Learning

Deep learning atau lebih dikenal dengan pembelajaran mendalam, adalah bagian dari metode *neural network* yang melakukan pemrosesan data secara tahapan pada lapisan atau *layer* (Goodfellow dkk., 2016). Terminologi dari *deep learning* didasari dari banyaknya atau dalamnya *layer* yang digunakan dalam membangun model. Konsep yang mendasari *deep learning* adalah *neural network*.

Deep learning memiliki basis arsitektur *neural network* yakni terdiri dari tiga jenis lapisan atau *layer*, yakni *input layer* sebagai lapisan masuknya data, *hidden layer* sebagai lapisan antara sekaligus tempat pemrosesan data, dan *output layer* sebagai lapisan dikeluarkannya hasil pemrosesan data (Azawii dkk., 2019). Setiap *layer* memiliki titik-titik yang menjadi tempat dilakukannya pemrosesan data yang disebut sebagai *node* atau *neuron*.



Gambar II.2 Visualisasi dari arsitektur *deep learning* (Azawii dkk., 2019)

Bagian *hidden layer* dalam hal ini dapat berjumlah lebih dari satu *layer*. Banyaknya *layer* yang digunakan disebut sebagai *depth* atau kedalaman dari model itu sendiri. Di sisi lain, banyaknya dimensi yang digunakan pada *hidden layer* disebut sebagai *width* atau lebar dari model itu sendiri. (Goodfellow dkk., 2016)

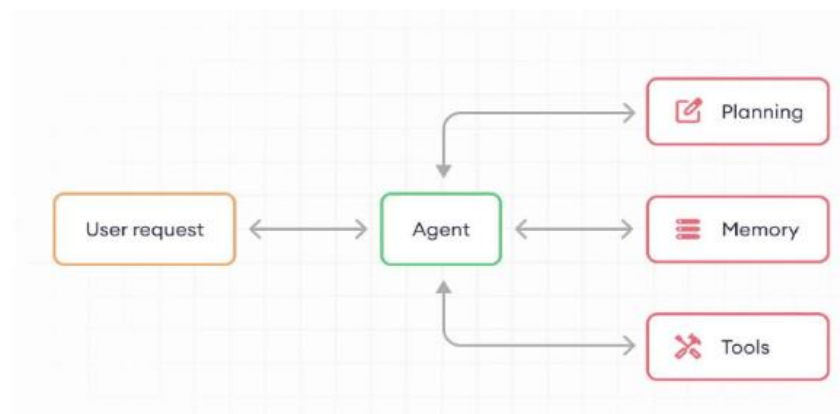
Berbasis pada *neural network*, seperti halnya jaringan saraf, model ini dinamakan *network* karena memiliki karakteristik untuk menggabungkan berbagai macam fungsi (f) sebagai satu kesatuan yang runtut. Maka dari itu, hasil *output* yang dikeluarkan oleh model ini dapat direpresentasikan sebagai untaian *input-output* fungsi yang terhubung berantai sebagai berikut:

$$y = f^n(f^{n-1}(\dots(f^2(f^1(x)))\dots)) \quad (\text{II-1})$$

dengan variabel n sebagai kedalaman dari model yang digunakan.

II.6 LLM Agents

LLM *agents* adalah sistem intelegensi buatan lanjut yang dikembangkan dan dirancang untuk menyelesaikan permasalahan yang membutuhkan urutan penalaran (*sequential reasoning*) (Bansal, 2024). Dalam penyelesaiannya, LLM *agents* melakukan pemecahan masalah menjadi permasalahan yang lebih kecil. Lalu, LLM *agents* memiliki kapabilitas untuk menggabungkannya kembali menjadi sebuah kesatuan.



Gambar II.3 Komponen sebuah LLM *agent* (Bansal, 2024)

Berdasarkan Gambar II.3, LLM *Agents* umumnya terdiri atas 4 komponen, yakni perencanaan (*planning*), ingatan (*memory*), kakas (*tools*), dan *agent* itu sendiri.

a. *Agent*

Inti dari LLM *agents* adalah *language models* yang dapat memproses dan memahami data yang diberikan kepadanya. Berdasarkan data yang diterimanya, LLM *agents* memiliki kapabilitas untuk mengambil keputusan secara otomatis. Dalam hal ini, LLM *agents* umumnya diberikan sebuah *prompt* spesifik. Maka dari itu, *prompt* memegang peran yang sangat penting karena memengaruhi bagaimana LLM *agents* tersebut akan merespon dan berinteraksi.

Pada komponen ini, persona dari *agent* juga dapat dikonfigurasi. Sebuah *agent* dapat diatur untuk memiliki sebuah karakteristik tertentu agar lebih cocok dalam menyelesaikan permasalahan tertentu. Hasil ini dapat dicapai dengan proses *tuning* pada *agent* itu sendiri.

b. *Memory*

Memory membantu LLM *agents* untuk menyelesaikan permasalahan yang sudah pernah diselesaikan sebelumnya. Berdasarkan jenisnya, *memory* dapat dibagi menjadi dua tipe, yakni *short-term memory* dan *long-term memory*. *Short-term memory* dibutuhkan oleh LLM *agents* untuk memahami konteks yang sedang berlangsung untuk memberikan respon sesuai konteks yang ada. Di sisi lain, *long-term memory* dibutuhkan agar LLM *agents* dapat belajar dari keseluruhan interaksi yang telah dilakukannya.

c. *Planning*

Planning digunakan oleh LLM *agents* untuk mendistribusikan permasalahan kompleks menjadi masalah yang lebih sederhana. Dalam hal ini, LLM *agents* perlu dapat beradaptasi pada berbagai permasalahan yang ada sehingga dihasilkan sebuah solusi yang optimal. Fase di dalam *planning* dibagi menjadi dua, yakni formulasi rencana (*plan formulation*) dan refleksi rencana (*plan reflection*).

Plan formulation adalah fase dimana *agents* mencoba untuk memecahkan masalah dengan membaginya menjadi masalah yang lebih sederhana. Beberapa metode yang umum digunakan dalam fase ini adalah *chain-of-*

thoughts (CoT) dan *tree-of-thoughts* (ToT). Di sisi lain, *plan reflection* adalah proses bagi *agent* untuk melakukan evaluasi atau *review* terkait penalaran yang sudah dilakukannya. Dengan begitu, *agents* berpotensi untuk memperbaiki kesalahan yang ada ataupun meningkatkan kualitas *output*.

d. *Tools*

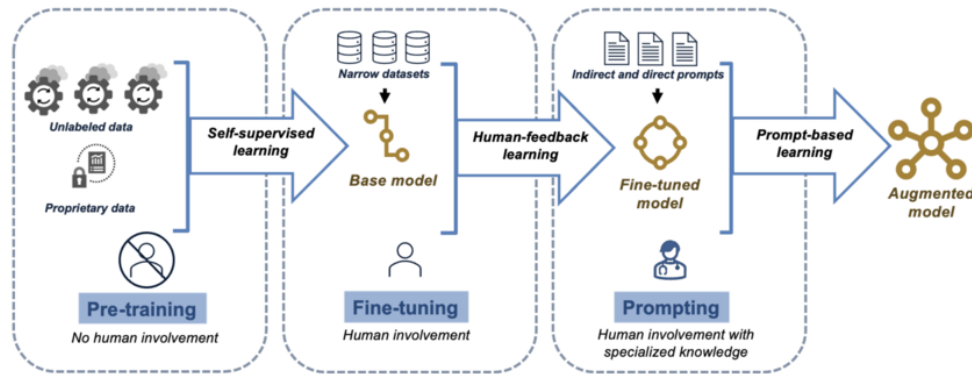
Tools didefinisikan sebagai berbagai sumber daya yang digunakan untuk membantu LLM *agents* terhubung dengan berbagai pihak eksternal dalam menyelesaikan permasalahan. Pekerjaan yang dilakukan oleh *tools* dalam hal ini umumnya mencakup, pengambilan data dari *database*, pembangkitan *query*, dan lain sebagainya. Setiap *tool* memiliki cara penggunaan *workflow* yang spesifik agar dapat digunakan oleh *agent*.

Proses pengambilan keputusan yang dilakukan oleh sistem LLM *agents* didasari oleh dengan keempat komponen yang tertera di dalamnya. Hal ini yang menjadi dasar bagi LLM *agents* dalam menyelesaikan berbagai permasalahan kompleks hingga bahkan melakukan proses pembelajaran secara mandiri. Terlebih lagi, tak jarang bagi sebuah LLM *agent* untuk melakukan sebuah pekerjaan secara kolaboratif dengan LLM *agent* lainnya. *Framework* tersebut dinamakan dengan *multi-agent LLMs*.

II.7 Large Language Models (LLMs)

Large Language Models (LLMs) merepresentasikan keahlian sistem komputasional yang mumpuni untuk memahami dan menghasilkan bahasa alami atau bahasa manusia (Parthasarathy dkk., 2024). LLMs, sebagai pengembangan dari *language models* (LMs) pada bidang pemrosesan bahasa alami (*natural language processing*), mampu menangani kasus seperti penanganan kata yang jarang digunakan (*rare word handling*), *overfitting*, dan penganan pola linguistik kompleks (*capturing complex linguistic patterns*). Konsep yang mendasari LLMs adalah penggunaan mekanisme *self-attention* pada arsitektur *transformer*.

Menurut Omiye (2023), proses dari LLMs dapat dibagi menjadi tiga bagian besar, yakni *pre-training*, *fine-tuning*, dan *prompting*. Secara gambaran besar, arsitektur LLMs dapat digambarkan dengan diagram berikut:



Gambar II.4 Arsitektur dari proses LLMs (Omiye dkk., 2023)

II.7.1 Knowledge Base (KB)

Knowledge base (KB) bisa diumpamakan sebagai sumber pengetahuan sekaligus tempat penyimpanan data dari sistem LLMs (Parthasarathy dkk., 2024). Umumnya, *knowledge base* dari sebuah sistem LLMs akan terus diperbarui agar sistem LLMs tetap mendapatkan pengetahuan dan informasi terbaru dan relevan. *Knowledge base* yang *outdated* pada LLMs dapat menyebabkan:

- Kesalahan fakta (*factual errors*) akibat informasi yang tidak diperbarui
- Irelevansi (*irrelevance*) akibat salahnya konteks yang ditangkap oleh sistem LLMs
- Kontinuitas bias (*bias perpetuation*) akibat kecondongan pada data yang terjadi karena tidak diperbaruinya data

II.7.2 Language Models

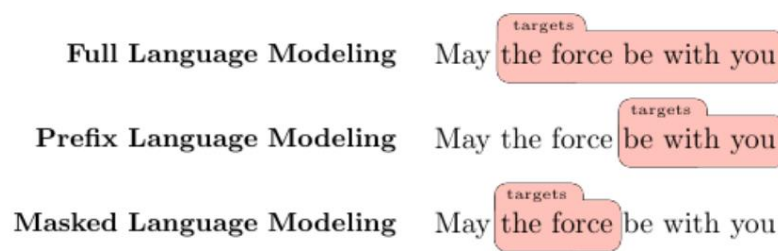
Menurut Coursera (2023), model dalam konteks *machine learning models* adalah program komputer yang memiliki kemampuan untuk mengenali pola dalam data dan melakukan prediksi. *Large language models* yang berfondasi pada *language models* merupakan bagian dari *machine learning models*. Proses untuk membentuk atau melatih *model* dinamakan sebagai *training*. Semakin tinggi intensitas *training*

yang diterapkan, semakin baik kualitas model yang akan dihasilkan. LLMs umumnya memiliki kemampuan untuk menyelesaikan berbagai pekerjaan pada tingkat performa seperti manusia dengan bayaran *training* yang perlu dilakukan pada model (Naveed dkk., 2024). *Training* yang dapat dilakukan mencakup: *parameter efficient tuning*, *pruning*, *quantization*, *knowledge distillation*, dan *context length interpolation*.

II.7.3 *Pre-trained Models*

Pre-trained model adalah model yang dibentuk dari proses *pre-training*. *Pre-trained model* seringkali juga disebut sebagai *base model* mengingat model ini adalah model dasar hasil pembelajaran pertama yang dilakukan pada sistem (Omiye dkk., 2023). *Pre-training* sendiri merupakan proses pembelajaran model secara *unsupervised* agar model memiliki pengetahuan dasar terkait data yang diberikan selama proses *pre-training*. Proses pelaksanaan *pre-training* dapat bervariasi bergantung pada pilihan *design* LLMs yang digunakan (Naveed dkk., 2024). *Pre-training* memiliki capaian (*objectives*) antara lain:

- a. *Full Language Modeling*, model auto regresi dimana pada prosesnya model diminta untuk melakukan prediksi terhadap *token* yang akan datang berdasarkan *token* yang telah diketahui
- b. *Prefix Language Modeling*, capaian *non-causal* dimana sebuah *prefix* dipilih secara acak dan hanya *target token* yang tersisa yang diperhitungkan
- c. *Masked Language Modeling*, capaian yang dicapai pada kondisi dimana sebuah *token* atau *sequence of token* disamarkan secara acak dan model diminta untuk melakukan prediksi terhadap *token* yang disamarkan tersebut sekaligus dengan *token* yang akan muncul kelak
- d. *Unified Language Modeling*, gabungan dari ketiga *modeling* yang telah disebutkan sebelumnya.



Gambar II.5 Visualisasi capaian dari *language model pre-training* (Naveed dkk., 2024)

Pre-trained model dapat dibagi-bagi berdasarkan *domain* kegunaannya. Namun, *general purpose model* menjadi model yang paling umum untuk digunakan. Berikut adalah beberapa contoh *pre-trained model* untuk *general purpose* (Naveed dkk., 2024):

- T5, *encoder-decoder model* untuk pembelajaran *unified text-to-text* dalam ranah permasalahan NLP, dengan menggunakan *masked language modelling* pada proses *pre-training*
- GPT, *transformer model* yang dilakukan proses pembelajaran atau pelatihan secara *batch*
- BLOOM, *causal decoder model* yang dilatih pada *ROOTS corpus* dan merupakan LLMs *open source*
- PaLM, *causal decoder* dengan mekanisme *attention* paralel dan lapisan *feedforward*
- LLaMA, *decoder-only model* yang terbuka pada operasi *parameter efficiency* dan *instruction tuning*

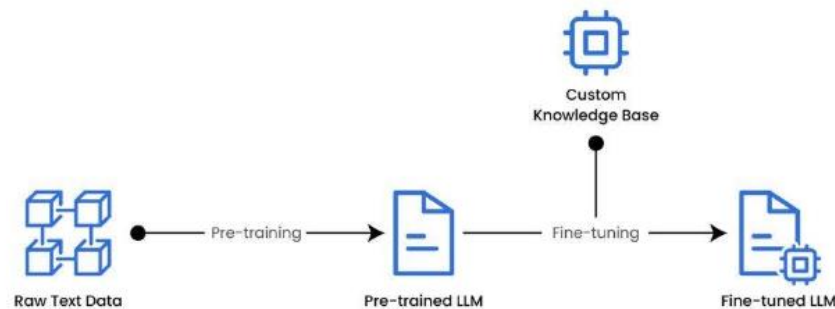
II.7.4 Teknik *Tuning*

Kapabilitas dari *pre-trained model* terbatas. Maka dari itu, pengembangan LLMs tidak berhenti sampai pada proses *pre-training*. Perlu dilakukan penyetelan (*tuning*) pada *pre-trained model* agar kualitas *pre-trained model* tetap terjaga. Ketika model tidak lagi di-*tuning* dan diperbarui, model dapat memberikan jawaban yang “mengarang” atas pernyataan yang tidak dapat mereka jawab. Fenomena ini sering

dikenal dengan *hallucination*. Dengan dilakukannya *tuning*, kualitas dan pemahaman model dapat dijaga atau bahkan ditingkatkan (Belcic & Stryker, 2024).

II.7.4.1 Fine Tuning

Fine-tuning adalah proses penyetelan model kepada sebuah *custom knowledge base* yang terkhusus (*specialized dataset*). Hal ini digunakan untuk melatih dan mempertajam pengetahuan model dalam *domain* tertentu.



Fine-tuning Process

Gambar II.6 Visualisasi proses pemodelan LLMs (Capella, 2024)

Menurut Naveed (2024), terdapat tiga pendekatan yang dapat dilakukan dalam mengimplementasikan *fine-tuning*. Pendekatan yang dilakukan adalah *transfer learning*, *instruction-tuning*, dan *alignment-tuning*.

a. *Transfer Learning*

Fine-tuning pada pendekatan ini dilakukan pada kasus bahwa *pre-trained LLMs model* telah bekerja dengan baik, tetapi ingin ditingkatkan pengetahuan model tersebut terhadap data tertentu (*task-specific data*).

b. *Instruction-tuning*,

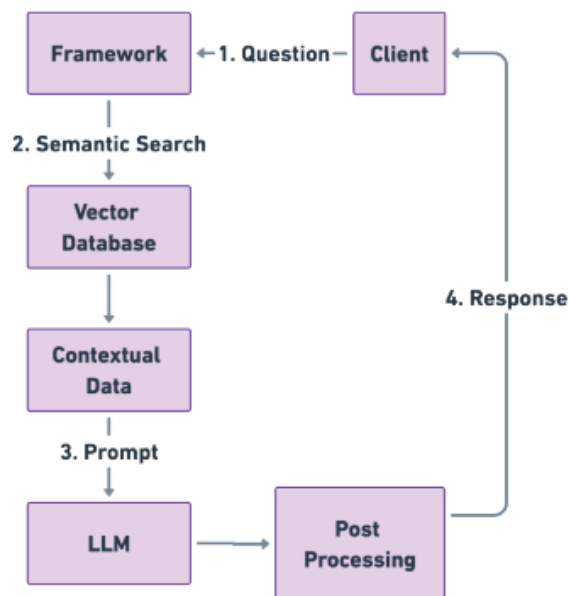
Fine-tuning pada pendekatan ini dilakukan pada *formatted data* yang terinstruksi, contohnya sebuah instruksi dan pasangan *input-output*. Instruksi dalam hal ini berguna untuk membimbing model untuk merespon berdasarkan *prompt* dan *input*.

c. *Alignment-tuning*

LLMs rentan untuk menghasilkan hasil yang salah ataupun bias. Maka dari itu, diperlukan cara untuk meluruskan atau mengoreksi hal ini menggunakan masukan dari manusia (*human feedback*)

II.7.4.2 Retrieval-Augmented Generation (RAG)

Retrieval-augmented generation (RAG) adalah proses substitusi yang dapat digunakan untuk menggantikan metode *fine-tuning*. Alih-alih bergantung pada pengetahuan di dalam *training data*, proses RAG membangun *workflows* untuk menarik informasi dan menghubungkan pada data *real-time* untuk menunjang performa dari model *LLMs* (Parthasarathy dkk., 2024).



Gambar II.7 Ilustrasi dari proses RAG (Parthasarathy dkk., 2024)

Proses dimulai dengan input *client* atau *user* berupa *query* kedalam sistem. Input dari *user* akan dilakukan pencarian menggunakan *semantic search* di dalam *vector database*. Proses dari *semantic search* ini menghasilkan data dengan konteks (*contextual data*) yang relevan dengan *query* yang dimasukkan oleh *user*. Setelah ditemukan, data tersebut dimasukkan sebagai *prompt* ke dalam model yang telah dibangun. *Output* dari model akan diproses terlebih dahulu sebelum akhirnya diberikan sebagai *response* kepada *user*.

II.7.5 Teknik *Prompting*

Prompting adalah metode untuk melatih LLMs untuk menghasilkan *output* (Naveed dkk., 2024). LLMs dapat di-*prompt* dengan berbagai cara. Terdapat kondisi dimana LLMs dapat beradaptasi dengan instruksi tanpa *fine-tuning* dan pada kondisi lain diperlukan *fine-tuning*. Hal ini berlaku karena adanya perbedaan dalam cara melakukan *prompt* (*prompt styles*). Penyetelan dari *prompt* ini merupakan bagian dari rekayasa *prompt* (*prompt engineering*) dimana *prompt* dibuat sedemikian rupa agar LLMs mampu mengerti *query* yang dimasukkan oleh pengguna. *Prompt* sendiri dapat memiliki setelan (*setups*) sebagai berikut:

- a. *Zero-Shot Prompting*, *style prompting* yang membutuhkan LLMs untuk menjawab pertanyaan pengguna tanpa melihat contoh atau jawaban pada *prompt*,
- b. *In-context Learning*, dikenal juga dengan *few-shot learning*, dalam *style* ini beberapa demonstrasi pasangan *input-output* dimasukkan untuk menghasilkan *response* yang sesuai,
- c. *Reasoning in LLMs*, *prompt* yang bertujuan untuk memprovokasi LLMs untuk menghasilkan jawaban untuk permasalahan logika,
- d. *Single-Turn Instructions*, dalam setelan ini, LLMs diekspektasikan untuk mengembalikan *response* berdasarkan pemahamannya dimana hanya terdapat satu kali interaksi dengan LLMs dan seluruh informasi relevan terdapat di dalam *prompt*, dan
- e. *Multi-Turn Instructions*, dalam setelan ini, dibutuhkan lebih dari satu interaksi dengan LLMs dimana siklus *feedback* dan *responses* digunakan sebagai *input* bagi LLMs.

Meski *prompt* berguna dalam proses di dalam LLMs, terdapat kondisi dimana pelaksanaan *prompt* secara manual dapat menurunkan performa dari LLMs (Naveed dkk., 2024). Maka dari itu, diperlukan sebuah *tuning* dalam pelaksanaan *prompting* itu sendiri, dinamakan *prompt tuning*.

Prompt tuning adalah ide untuk melakukan *tuning* pada *prompt* yang berkelanjutan (Liu dkk., 2022). Dalam pelaksanaannya, hanya *prompt* yang berkelanjutanlah yang

diperbarui selama proses *training*. Tujuan dari *prompt tuning* sendiri adalah untuk meningkatkan efisiensi pembelajaran dari model dengan mengurangi biaya waktu dan biaya penyimpanan (*time memory and per-task storage cost*). Dengan dilakukannya *prompt tuning*, proses pembelajaran dapat dilakukan dengan akurasi yang lebih tinggi (*high accuracy*) dan penggunaan parameter yang lebih efisien (*parameter efficiency*), menjadikan proses *prompt tuning* sebagai alternatif dari proses *fine-tuning* (Liu dkk., 2022).

II.7.6 State Memory Flows

State Memory Flows adalah alur bagi LLMs untuk tetap memahami konteks yang sesuai. Dalam sebuah percakapan, terdapat banyak topik dan konteks yang diberikan oleh pengguna kepada LLMs. Maka dari itu, dibutuhkan pengelolaan terkait ingatan LLMs terkait konteks yang sesuai (*contextual memory management*) (Porcu, 2024). Menurut Porcu, terdapat 4 cara yang digunakan untuk mengelola *contextual memory*, yaitu:

- a. *Time-based decay*, proses pengelolaan konteks dimana relevansi dari konteks diukur berdasarkan umur dari konteks tersebut
- b. *Event-driven forgetting*, proses pengelolaan konteks berdasarkan *event* tertentu
- c. *Memory compression*, proses pengelolaan konteks dengan pemampatan *storage* yang lebih efisien
- d. *Dynamic memory allocation*, proses pengelolaan konteks dengan alokasi *memory* berdasarkan kompleksitas *query*

II.7.7 Self-Detection

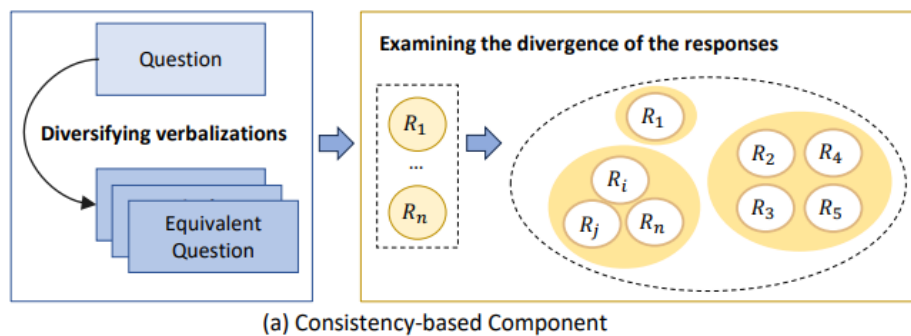
Sistem LLM pada umumnya masih memiliki keterbatasan dalam memahami informasi (Yin dkk., 2023). Maka dari itu, kemampuan untuk mengetahui batasan dari sebuah LLM adalah hal yang penting. Kemampuan untuk mengenali batasan ini, atau lebih dikenal dengan “*knowing what you don’t know*“, dapat berguna untuk meningkatkan performa dari LLM itu sendiri. Dengan begitu, sistem LLM dapat mengenali kemampuannya bila diperhadapkan dengan pertanyaan yang diluar

lingkup pengetahuannya atau bahkan pertanyaan yang memang tidak ada jawabannya. Konsep untuk mengenali batasan ini, pada LLM lebih umum dikenal sebagai *self-detection*.

Menurut Zhao dkk. (2024), terdapat dua *framework* yang dapat digunakan untuk menerapkan konsep *self-detection*, yakni *consistency-based component* dan *verbalization-based component*.

a. *Consistency-based component*

Proses dari *consistency-based component* dapat dibagi menjadi dua bagian, yakni bagian pemecahan pertanyaan (*diversifying question*) dan perhitungan konsistensi (*consistency score*). Visualisasi dari komponen ini dapat ditampilkan sebagai berikut:



Gambar II.8 Visualisasi *consistency-based component* (Zhao dkk., 2024)

Proses pemecahan pertanyaan dilakukan dengan membagi suatu pertanyaan menjadi beberapa pertanyaan serupa. Proses ini sendiri bisa dilakukan dengan dua cara, yakni *model-based generation* dan *rule-based generation*. *Model-based generation* dilakukan dengan meminta sebuah sistem LLM untuk menghasilkan pertanyaan serupa. Di sisi lain, *rule-based generation* menggunakan sebuah aturan tertentu yang telah ditetapkan untuk menghasilkan pertanyaan.

Jawaban atas pertanyaan-pertanyaan tersebut akan dihitung nilai konsistensinya. Semakin tinggi nilai konsistensi merepresentasikan

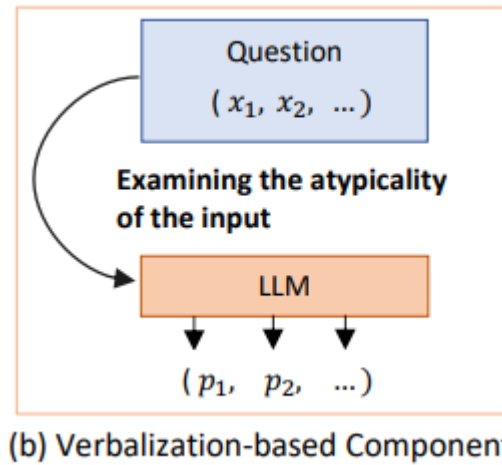
semakin tinggi keyakinan model dalam menjawab pertanyaan tersebut. Di sisi lain, rendahnya konsistensi menunjukkan ketidakpahaman model atas pertanyaan yang diberikan.

b. *Verbalization-based component*

Verbalization-based component melakukan proses perhitungan *score* menggunakan *negative log-likelihood* pada *verbalized input*. Persamaan yang digunakan antara lain:

$$A(q) = -\log P(q) = - \sum_t^T \log P(x_t | X_{<t}) \quad (\text{II-2})$$

Variabel x_t dan $X_{<t}$ menunjukkan *token* dan *token set* dari pertanyaan q . Setelah nilai $A(q)$ diperoleh, dilakukan perhitungan nilai normalisasi $A(q)/N(q)$, dengan $N(q)$ adalah jumlah *token* pada pertanyaan q . Nilai $A(q)$ yang semakin tinggi menunjukkan bahwa hasil yang didapatkan mengarah kearah ‘ketidakselarasan’ (*atypicality*). Berikut adalah visualisasi dari *verbalization-based component*.



Gambar II.9 Visualisasi *verbalization-based component* (Zhao dkk., 2024)

II.8 Model Perilaku Pengguna

Di dalam media sosial, sebuah data yang menggambarkan jaringan dan interaksi pengguna dapat diperoleh. Data ini dinamakan sebagai *social network data*. *Social network data* mengandung 2 elemen utama (Jiang & Ferrara, 2023) yaitu:

- a. konten (*content*) yang berisikan isi hal-hal yang dibagikan oleh pengguna
- b. jaringan (*network*) yang merepresentasikan dengan siapa, kapan, dan bagaimana interaksi antar pengguna dilakukan.

Pemodelan dari *social network data* umumnya berbasis pada representasi graf dan berlandaskan konsep *social network homophily*. Konsep *network homophily* adalah konsep pada pembelajaran jaringan yang menyatakan *node* yang memiliki hubungan umumnya memiliki kemiripan. Dalam konteks *social network homophily*, dapat diartikan sebagai pengguna yang saling terhubung cenderung untuk memiliki kemiripan.

Sistem LLMs dapat digunakan untuk melakukan pemrosesan terkait *homophily*. Aspek yang diproses oleh sistem LLMs adalah *linguistic homophily* di dalam *social network data* sehingga pengguna dengan kemiripan gaya bahasa memiliki kecenderungan untuk berteman. Sesuai dengan elemen yang ada pada *social network data*, pembelajaran data ini dilakukan berdasarkan dua fitur, yaitu konten (*content cues*) dan jaringan (*network cues*).

Pada fitur konten, data yang diperoleh adalah konten *textual* seperti deskripsi profil pengguna. Deskripsi profil ini dapat digunakan selayaknya biografi kecil yang merepresentasikan sang pengguna itu sendiri. Selain deskripsi profil, dapat digunakan metadata dari akun pengguna itu sendiri seperti jumlah pengikut dan tanggal penciptaan akun serta *post* dari pengguna itu sendiri. Di sisi lain, dalam fitur jaringan, data yang diperoleh mencakup interaksi yang dilakukan oleh pengguna, yakni: *follow*, *share*, dan *like*. Cara dan kepada siapa pengguna melakukan interaksi di dalam sosial media dapat menggambarkan perilaku dari sang pengguna. Maka dari itu, proses pembelajaran dari *social network data* bertujuan untuk mempelajari model perilaku pengguna (*user behavior model*).

BAB III

ANALISIS MASALAH DAN RANCANGAN SOLUSI

PENGEMBANGAN LLM PADA *SOCIAL MEDIA BOT*

INSTAGRAM

Dalam menyelesaikan tugas akhir, dibutuhkan pemahaman yang mendalam untuk menentukan solusi optimal dalam menyelesaikan permasalahan yang telah ditentukan. Maka dari itu, dibutuhkan analisis terkait masalah untuk masalah umum persoalan *capstone* dan masalah spesifik dalam pelaksanaan tugas akhir ini. Tak hanya itu, dilakukan juga analisis terkait solusi maupun alternatif solusi yang dapat diimplementasikan untuk menyelesaikan permasalahan tersebut.

III.1 Deskripsi Umum Persoalan *Capstone*

Persoalan yang ditanggulangi dalam pengerjaan Tugas Akhir *Capstone* ini adalah pengembangan *social media bot* pada media sosial Instagram yang bertindak selayaknya seorang *influencer* pariwisata. *Social media bot* ini dilengkapi dengan implementasi LLMs agar dapat berinteraksi secara alami dengan pengguna dalam bahasa yang dimengerti oleh manusia. Alih-alih mendekatkan manusia untuk mengerti bahasa yang digunakan oleh mesin, penggunaan LLMs pada *social media bot* menggunakan pendekatan sebaliknya, yakni mendekatkan mesin untuk mengerti bahasa alami manusia.

III.2 Analisis Masalah

Pariwisata adalah sektor yang berkembang pesat di Indonesia. Hal ini direpresentasikan dengan sektor pariwisata yang meningkat setiap tahunnya. Namun, pembangunan ini masih belum merata, Masih banyak tempat wisata di Indonesia yang, meski memiliki potensi besar, masih belum diketahui oleh banyak turis mancanegara. Salah satu contohnya adalah tempat wisata di Nusa Tenggara Barat dan Nusa Tenggara Timur.

Meski *social media bot* sudah sering diterapkan dalam berbagai media sosial, *social media bot* pada umumnya masih jauh dari imitasi perilaku manusia. Kecenderungan dari *bot* tersebut adalah memberikan respons yang pasif dan kaku. Kemampuan interaksi *bot* yang terbatas tentunya tidak cukup bila digunakan untuk mendorong keinginan para pengguna. Maka dari itu, kualitas dari *social media bot* perlu ditingkatkan dengan implementasi LLMs sehingga *social media bot* mampu berinteraksi secara alami selayaknya seorang manusia.

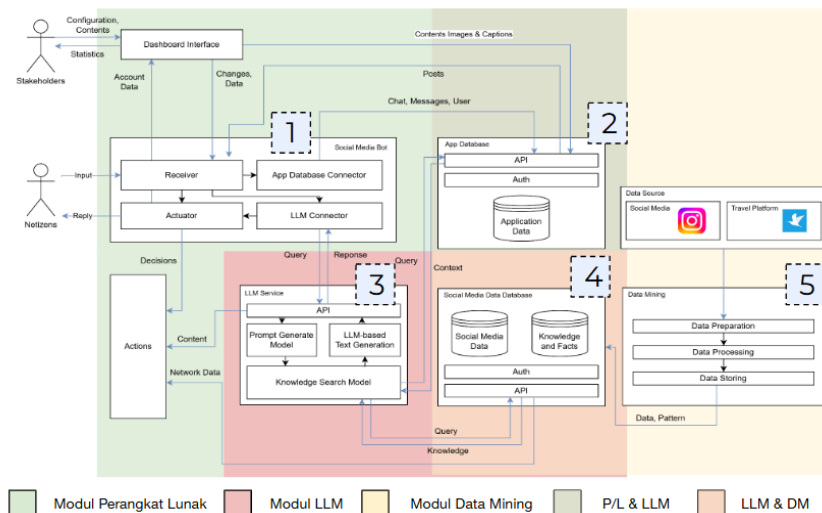
III.3 Analisis Solusi

Social media bot ini perlu memiliki kapabilitas untuk mendorong para pengguna Instagram untuk melakukan perjalanan wisata dengan berlaku sebagai seorang *influencer*. Dalam implementasinya, *bot* ini dilengkapi dengan sistem LLMs agar *bot* dapat memahami perilaku manusia dan memberikan *output* selayaknya manusia. Maka dari itu, rincian solusi yang akan dilakukan dalam pengembangan sistem LLMs dapat dijelaskan sebagai berikut:

1. *Pretrained model* yang digunakan sebagai basis dari sistem LLMs adalah model LLaMA 3. Model ini dipilih karena memiliki kecocokan dalam pengembangan LLM untuk *domain* spesifik dan penggunaannya yang *open-source*.
2. Teknik *tuning* yang digunakan adalah *retrieval-augmented generation* (RAG) dengan alasan kecocokan dengan *dataset* yang tidak terlalu besar dan kebutuhan akan *real-time data*. Dalam hal ini, Teknik *fine-tuning* dipilih sebagai alternatif solusi dalam meningkatkan performa model.
3. Teknik *prompting* yang digunakan adalah *few-shot learning*, *chain of thoughts* dengan alasan kecocokan dengan teknik RAG dan kebutuhan untuk proses penalaran. Selanjutnya, jenis interaksi yang digunakan adalah *multi-turn instructions* yang membutuhkan interaksi dua arah antara LLM dan *user*.
4. Implementasi kustomisasi yakni, *state memory flows* agar LLM tetap mengetahui konteks dari percakapan dan *self-detection* untuk mencegah LLM dalam mengalami *hallucination*.

III.4 Deskripsi Solusi

Berdasarkan analisis solusi yang telah dipaparkan, dikembangkan sebuah *social media bot*. Bot ini penulis beri nama ‘Nusava’ untuk menekankan persona *influencer* dari bot sesuai penyelesaian permasalahan yang telah dibahas sebelumnya.



Gambar III.1 Diagram arsitektur dari *social media bot* Nusava

Sistem yang telah dirancang akan dikembangkan menggunakan bahasa pemrograman Python. Komponen dari *social media bot* yang dikembangkan dalam hal ini mencakup lima komponen utama yakni:

1. Sistem *social media bot* itu sendiri, yang berfungsi sebagai sistem utama sekaligus *interface* interaksi dengan pengguna
2. *App database* yang digunakan untuk menyimpan *context* dan *session* yang telah dilakukan oleh *user* dengan *bot*
3. Sistem LLM yang berfungsi selayaknya otak pada *social media bot* agar *bot* mampu memiliki pemahaman dan berperilaku selayaknya manusia
4. *Knowledge database* yang digunakan untuk menyimpan *knowledge* dan data terkait pariwisata
5. Sistem *data mining* yang berfungsi sebagai sumber *knowledge* dari *bot*

Keseluruhan rancangan yang telah dipaparkan pada Bagian III.3 akan berada di dalam komponen nomor 3, yakni sistem LLM. Dalam penerapannya kelak, sistem LLM akan saling berinteraksi dengan komponen lain.

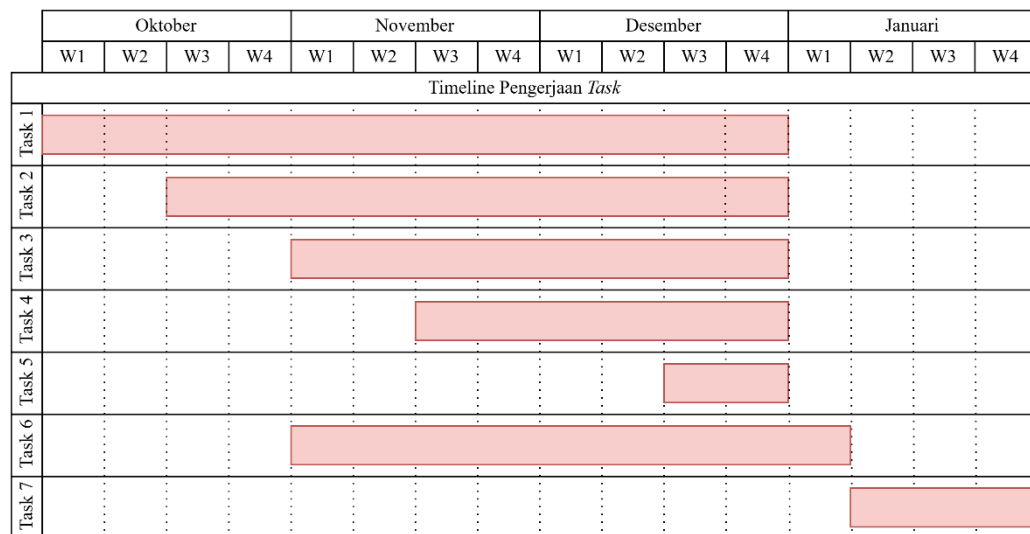
BAB IV

RENCANA PELAKSANAAN

Pelaksanaan tugas akhir membutuhkan perencanaan yang matang. Setiap tahapan proses dalam pengerjaan tugas akhir perlu direncanakan dalam sebuah jadwal yang konkrit dan spesifik. Di sisi lain, berbagai risiko dapat muncul sebagai hambatan dalam pengerjaan tugas akhir. Maka dari itu, dibutuhkan perancangan rencana mitigasi yang dapat digunakan, baik sebagai langkah pencegahan maupun penanggulangan, untuk risiko yang mungkin terjadi selama pengerjaan tugas akhir.

IV.1 Jadwal

Jadwal pengerjaan tugas akhir dimulai dengan dilaksanakannya penyusunan proposal. Linimasa bagi pelaksanaan penyusunan proposal dapat digambarkan dengan *gantt chart* sebagai berikut:



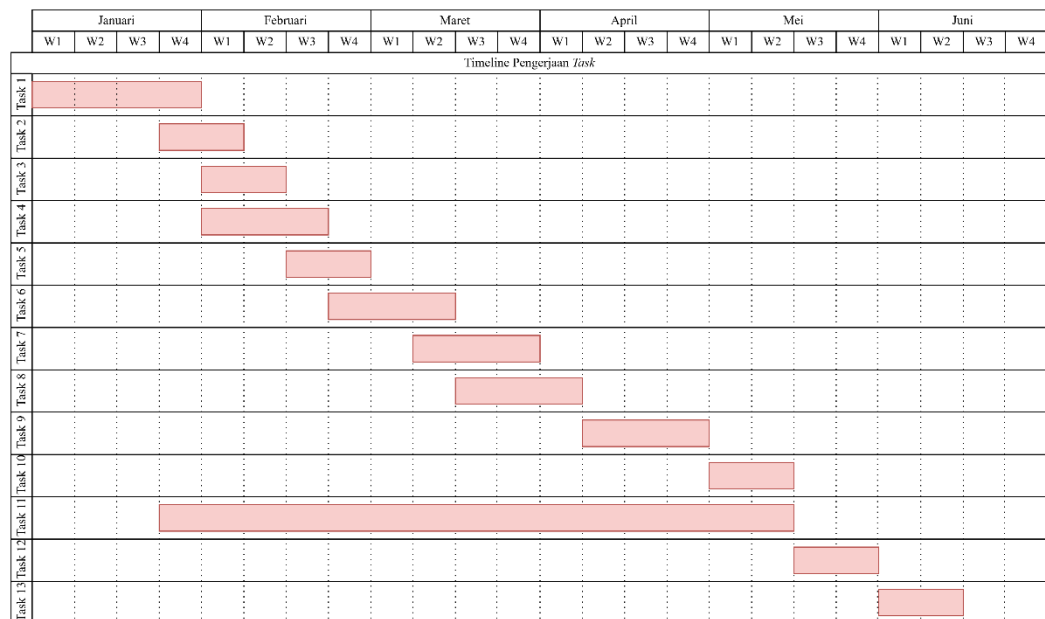
Gambar IV.1 *Gantt chart* pengerjaan proposal tugas akhir

Berdasarkan Gambar IV.1, terdapat tujuh *task* yang telah diselesaikan selama pengerjaan proposal tugas akhir ini. Penjelasan dari tujuh *task* tersebut antara lain:

Tabel IV.1 Deskripsi *task* dalam pengerjaan proposal tugas akhir

| No | Deskripsi <i>task</i> |
|----|--|
| 1 | Pelaksanaan studi literatur |
| 2 | Penulisan Bab 2 proposal tugas akhir |
| 3 | Penulisan Bab 1 proposal tugas akhir |
| 4 | Penulisan Bab 3 proposal tugas akhir |
| 5 | Penulisan Bab 4 proposal tugas akhir |
| 6 | Pelaksanaan revisi proposal tugas akhir |
| 7 | Pelaksanaan seminar proposal tugas akhir |

Di sisi lain, jadwal dari pengerjaan tugas akhir ini akan dimulai dari Bulan Januari hingga Bulan Juni tahun 2025. Rencana dari pengerjaan tugas akhir dapat digambarkan dengan *gantt chart* sebagai berikut:



Gambar IV.2 *Gantt chart* pengerjaan tugas akhir

Berdasarkan Gambar IV.2, terdapat 13 *task* yang perlu diselesaikan selama pengerjaan tugas akhir ini. Berikut adalah penjelasan dari 13 sistem yang tertera pada *gantt chart* tersebut:

Tabel IV.2 Deskripsi *task* dalam pengerjaan tugas akhir

| No | Deskripsi <i>task</i> |
|----|---|
| 1 | Pelaksanaan eksplorasi terkait implementasi yang akan dilakukan |
| 2 | Pelaksanaan implementasi penggunaan <i>pre-trained model</i> pada sistem |
| 3 | Pelaksanaan integrasi dengan <i>database</i> Modul <i>data mining</i> |
| 4 | Pelaksanaan implementasi sistem <i>retrieval-augmented generation</i> (RAG) |
| 5 | Pelaksanaan integrasi dengan <i>database</i> Modul perangkat lunak |
| 6 | Pelaksanaan implementasi mekanisme <i>prompting</i> |
| 7 | Pelaksanaan implementasi <i>state memory flows</i> |
| 8 | Pelaksanaan implementasi <i>self-detection</i> |
| 9 | Pelaksanaan <i>tuning</i> dan <i>improvement</i> terhadap kinerja model |
| 10 | Pelaksanaan <i>deployment</i> dan implementasi API |
| 11 | Penulisan laporan tugas akhir |
| 12 | Pelaksanaan seminar tugas akhir |
| 13 | Pelaksanaan sidang tugas akhir |

IV.2 Risiko

Risiko adalah hal yang mungkin terjadi selama pengerjaan tugas akhir ini. Maka dari itu, perlu direncanakan langkah-langkah yang dapat dilakukan untuk mencegah maupun menanggulangi risiko tersebut. Berikut adalah risiko-risiko yang mungkin terjadi selama pengerjaan tugas akhir:

Tabel IV.3 Daftar risiko dan rencana mitigasi dari risiko

| No | Risiko | Rencana Mitigasi |
|----|--|--|
| 1 | Sistem <i>bot</i> tidak berjalan pada <i>platform</i> Instagram karena terdeteksi kecurigaan sebagai <i>automated system</i> | Menyiapkan akun Instagram cadangan dan membuat aksi <i>bot</i> memiliki <i>delay</i> dalam rentang waktu yang acak |

| | | |
|---|---|--|
| 2 | Kesibukan penulis dalam mengerjakan berbagai tugas selain tugas akhir | Meluangkan waktu tertentu secara khusus untuk melanjutkan progres pengerjaan tugas akhir |
| 3 | Jadwal pelaksanaan tugas akhir yang tidak sesuai dengan perencanaan | Membuat <i>priority list</i> dalam setiap pengerjaan tugas dan jadwal penyesuaian agar penyelesaian <i>task</i> tetap sesuai jadwal mula-mula |
| 4 | Pemilihan solusi tidak tepat atau tidak sesuai dengan kebutuhan | Dilakukannya eksplorasi tambahan terkait alternatif-alternatif yang dapat dilakukan dalam menyelesaikan permasalahan |
| 5 | Penerapan sistem LLMs yang masih tidak cukup untuk mengimitasi perilaku <i>influencer</i> | Penambahan algoritma tambahan dengan melakukan pengamatan secara manual terhadap perilaku <i>influencer</i> sehingga dapat diterapkan kepada sistem <i>bot</i> |

DAFTAR PUSTAKA

Azawii, A., Al-Janabi, S. T. F., & Al-Khateeb, B. (2019). *Survey on Intrusion Detection Systems based on Deep Learning*.
<https://doi.org/10.21533/pen.v7i3.635>

Badan Pusat Statistik. (2023). *The tourism industry continues to recover, as shown by the 52.76 percent rise in international visitor arrivals in September 2023 over the same month last year*.
<https://www.bps.go.id/en/pressrelease/2023/11/01/2047/the-tourism-industry-continues-to-recover--as-shown-by-the-52-76-percent-rise-in-international-visitor-arrivals-in-september-2023-over-the-same-month-last-year.html>

Badan Pusat Statistik. (2024). *Jumlah Kunjungan Wisatawan Mancanegara ke Indonesia Menurut Kebangsaan (Kunjungan), 2021-2023 [Dataset]*.
<https://www.bps.go.id/id/statistics-table/2/MTgyMSMy/jumlah-kunjungan-wisatawan-mancanegara-ke-indonesia-menurut-kebangsaan.html>

Bansal, A. (2024). *Comprehensive Study on LLM Agents*.
https://www.researchgate.net/publication/386093988_Comprehensive_Study_on_LLM_Agents

Belcic, I., & Stryker, C. (2024, Agustus). *RAG vs. Fine-tuning*.
<https://www.ibm.com/id-id/think/topics/rag-vs-fine-tuning>

Capella. (2024, Juli 7). *Strategies for Fine-Tuning Large Language Models*.
<https://www.capellasolutions.com/blog/strategies-for-fine-tuning-large-language-models>

Chakraborty, P. (2024). *Social Media and Tourism: Key Statistics for 2024*. Winsavvy. <https://www.winsavvy.com/social-media-and-tourism-key-statistics/#:~:text=Hashtags%20on%20Instagram%3F-Travel%2DRelated%20Hashtags%20on%20Instagram%20Have%20Over%201%20Billion%20Posts,experiences%20and%20discover%20new%20destinations.>

Coursera. (2023, November 30). *Machine Learning Models: What They Are and How to Build Them*. <https://www.coursera.org/articles/machine-learning-models>

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. <https://www.deeplearningbook.org/>

Grimme, C., Preuss, M., Clever, L., & Trautmann, H. (2017). *Social Bots: Human-Like by Means of Human Control?* <https://doi.org/10.1089/big.2017.0044>

Indusface. (2024). *Social Media Bots – What They Are, Why They’re Used & How to Prevent Them*. <https://www.indusface.com/learning/social-media-bots/>

Jiang, J., & Ferrara, E. (2023). *Social-LLM: Modeling User Behavior at Scale using Language Models and Social Network Data*. <https://arxiv.org/pdf/2401.00893>

- Liu, X., Ji, K., Fu, Y., Tam, W. L., Du, Z., Yang, Z., & Tang, J. (2022). *P-Tuning v2: Prompt Tuning Can Be Comparable to Fine-tuning Universally Across Scales and Tasks*. 8.
- Magnus, T., & Hakan, S. (2022). *Act like a human, think like a bot*. 56.
- Naveed, H., Khan, A. U., & Qiu, S. (2024). *A Comprehensive Overview of Large Language Models*. 47.
- Oestreicher, G. (2024, November 18). *Social Media Bots: The Good and The Bad*. <https://metricool.com/social-media-bots/>
- Omiye, J. A., Gui, H., & Rezaei, S. J. (2023). *Large language models in medicine: The potentials and pitfalls*. https://www.researchgate.net/figure/Overview-of-LLM-training-process-LLMs-learn-from-more-focused-inputs-at-each-stage-of_fig1_373642018
- Parthasarathy, V. B., Zafar, A., Khan, A., & Shahid, A. (2024). *The Ultimate Guide to Fine-Tuning LLMs from Basics to Breakthroughs: An Exhaustive Review of Technologies, Research, Best Practices, Applied Research Challenges and Opportunities*. <https://arxiv.org/pdf/2408.13296>
- Porcu, V. (2024). *The Role of Memory in LLMs: Persistent Context for Smarter Conversations*. 19.
- Rusyidi, B., & Fedryansah, M. (2018). *PENGEMBANGAN PARIWISATA BERBASIS MASYARAKAT. 1*, 11.
- Setiawan, I. (2015). *POTENSI DESTINASI WISATA DI INDONESIA MENUJU KEMANDIRIAN EKONOMI*.

<https://media.neliti.com/media/publications/173034-ID-potensi-destinasi-wisata-di-indonesia-me.pdf>

We Are Social. (2024). *Digital 2024 Indonesia* (hlm. 136).
<https://indd.adobe.com/view/99d51a14-cbfe-48d4-bb25-ace9496ee758?allowFullscreen=true>

Wibowo, M. S., & Belia, L. A. (2023). Partisipasi Masyarakat dalam Pengembangan Pariwisata Berkelanjutan. *JURNAL MANAJEMEN PERHOTELAN DAN PARIWISATA*, 6(1), 8.

Widada, C. K. (2018). *MENGAMBIL MANFAAT MEDIA SOSIAL DALAM PENGEMBANGAN LAYANAN*. <http://dx.doi.org/10.33505/jodis.v2i1.130>

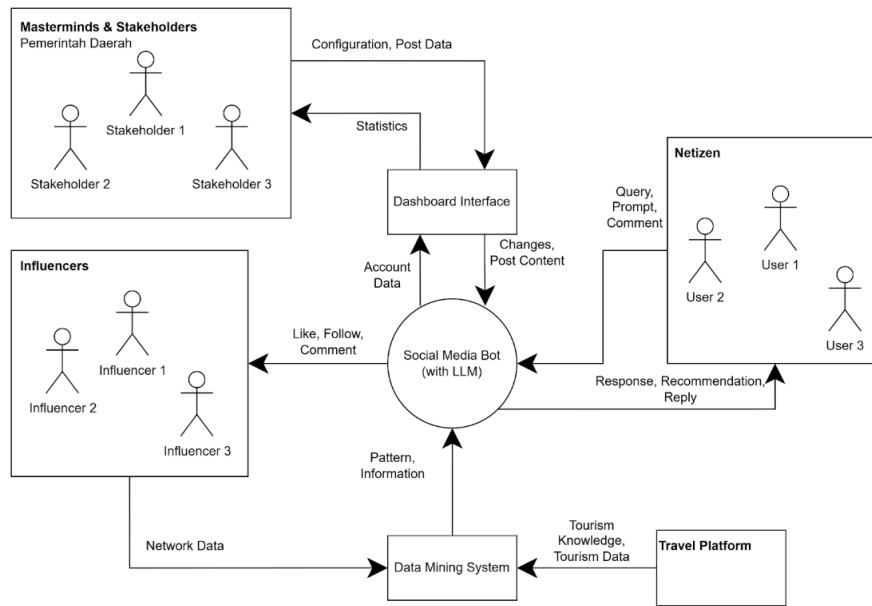
Yin, Z., Sun, Q., & Guo, Q. (2023). *Do Large Language Models Know What They Don't Know?* 10.

Zhao, Y., Yan, L., Sun, W., Xing, G., & Meng, C. (2024). *Knowing What LLMs DO NOT Know: A Simple Yet Effective Self-Detection Method*. 13.

Zote, J. (2024, Februari 22). *Instagram statistics you need to know for 2024 [Updated]*. <https://sproutsocial.com/insights/instagram-stats/>

LAMPIRAN A. RENCANA UMUM PROYEK

Proyek dari Tugas Akhir ISML-03 adalah proyek pengembangan *social media bot* pada *platform* Instagram. Maka dari itu, setelah peluncurannya, *bot* akan berinteraksi dengan berbagai pihak dalam keberjalanannya. Setiap pihak akan memiliki interaksi yang berbeda dengan *bot* menyesuaikan dengan kebutuhan yang juga berbeda. Berikut adalah diagram posisi *bot* pada media sosial:



Gambar A.1 Diagram posisi *social media bot*

Terdapat beberapa pihak yang akan memiliki interaksi langsung dengan *social media bot* yang dikembangkan. Pihak-pihak tersebut dikelompokkan menjadi tiga kelompok antara lain: *netizen*, *stakeholders*, dan *influencers*. Keterangan dari interaksi setiap pihak dijelaskan secara rinci sebagai berikut:

Tabel A.1 Daftar pihak dan interaksinya dengan *bot*

| Pihak | Kebutuhan <i>Bot</i> terhadap Pihak | Kebutuhan Pihak terhadap <i>Bot</i> | Media Interaksi |
|-------|-------------------------------------|-------------------------------------|-----------------|
| | | | |

| | | | |
|---------------------|--|--|--|
| <i>Stakeholders</i> | Pemenuhan data dan konten yang perlu diunggah oleh <i>bot</i> | Statistik dan data terkait performa dan aksi yang telah dilakukan oleh bot | <i>Dashboard interface</i> |
| <i>Netizen</i> | Masukan dari pihak berupa pertanyaan dalam fitur <i>direct messages</i> dan tanggapan dalam kolom komentar | Informasi dan rekomendasi terkait pariwisata | Fitur <i>direct messages</i> (DM), <i>post</i> milik <i>bot</i> , dan kolom komentar dari <i>post</i> milik <i>bot</i> |
| <i>Influencers</i> | <i>Post</i> dan koneksi dari akun <i>influencers</i> | - | <i>Post</i> para <i>influencers</i> dan kolom komentar dari <i>post</i> para <i>influencers</i> |

Selama pengerjaannya, pengembangan social media bot akan dibagi menjadi tiga modul yang berbeda, yakni Modul Sistem Otomasi, Modul *Large Language Model* (LLM), dan Modul Penambangan Data. Setiap modul memiliki deskripsi fungsional sebagai berikut:

1. Modul Sistem Otomasi bertanggung jawab untuk mengembangkan sistem *social media bot* sehingga dapat diintegrasikan dengan media sosial Instagram, Modul LLM, dan Modul Penambangan Data, serta memiliki kapabilitas untuk melakukan proses input-output dengan pengguna.
2. Modul LLM bertanggung jawab untuk memproses *input* yang diterima oleh Modul Perangkat Lunak dengan basis intelegensi buatan dan pembelajaran mesin sehingga dihasilkan *output* yang menyerupai tingkat kecerdasan manusia berlandaskan pengetahuan dari Modul Penambangan Data.

3. Modul Penambangan Data bertanggung jawab untuk melakukan proses perolehan pengetahuan dan pola dengan memenuhi kebutuhan data, untuk Modul Perangkat Lunak dan Modul LLM, baik dari sumber internal sosial media maupun sumber eksternal.

LAMPIRAN B. SPESIFIKASI KEBUTUHAN SISTEM

Lampiran ini berisikan spesifikasi lengkap terkait kebutuhan fungsional dan non-fungsional dari sistem *social media bot* yang dikembangkan. Berikut adalah spesifikasi dari kebutuhan fungsional sistem:

Tabel B.1 Spesifikasi kebutuhan fungsional sistem

| ID | Modul | Kebutuhan |
|-----|------------------|---|
| F01 | Sistem Otomasi | Modul dapat menerima <i>Direct Message</i> (DM) dari pengguna lain dan mengirimkannya ke Modul LLM. |
| F02 | Sistem Otomasi | Modul dapat memberikan balasan otomatis kepada pengguna lain berdasarkan <i>output</i> yang diberikan oleh Modul LLM. |
| F03 | Sistem Otomasi | Modul dapat melakukan <i>like</i> , <i>comment</i> , dan <i>follow</i> pada akun pengguna lain yang telah ditentukan oleh Modul Penambangan Data secara otomatis dan berkala. |
| F04 | Sistem Otomasi | Modul dapat membalas komentar yang diterima pada postingan pengguna secara otomatis. |
| F05 | <i>Dashboard</i> | Modul menyediakan antarmuka untuk melakukan konfigurasi profil akun pengguna. |
| F06 | <i>Dashboard</i> | Modul dapat menampilkan <i>log</i> aktivitas <i>bot</i> serta <i>insight</i> dari seluruh postingan pengguna. |
| F07 | <i>Dashboard</i> | Modul memungkinkan pengguna untuk mengunggah gambar dan menambahkan <i>caption</i> yang akan digunakan untuk membuat postingan baru. |
| F08 | LLM | Modul perlu memiliki kapabilitas untuk memahami <i>input</i> pengguna dalam bahasa alami. |
| F09 | LLM | Modul dapat memberikan <i>output</i> kepada pengguna dalam bahasa alami. |
| F10 | LLM | Modul terhubung dengan sistem <i>database</i> dan mengeluarkan hasil yang dipengaruhi oleh data atau informasi pada <i>database</i> . |

| | | |
|-----|------------------|---|
| F11 | LLM | Modul dapat melakukan pencarian informasi relevan ke <i>database</i> yang disediakan oleh Modul Penambangan Data. |
| F12 | LLM | Modul dapat menghasilkan <i>prompt</i> yang sesuai dengan kebutuhan. |
| F13 | LLM | Modul dapat menangani pergantian <i>state</i> dan konteks. |
| F14 | LLM | Modul perlu mengetahui batas ketidaktahuannya dan tidak menghasilkan hasil <i>hallucination</i> . |
| F15 | LLM | Modul dapat diintegrasikan dan digunakan oleh Modul Perangkat Lunak. |
| F16 | Penambangan Data | Modul dapat melakukan <i>scraping</i> terhadap data - data di <i>travel website</i> dan menyimpannya di dalam <i>database</i> |
| F17 | Penambangan Data | Modul dapat mendapatkan <i>frequent pattern</i> dari informasi - informasi umum yang sudah tersimpan |
| F18 | Penambangan Data | Modul dapat membentuk <i>social network</i> dari pengguna - pengguna Instagram yang berkaitan dengan pariwisata Nusa Tenggara dan menyimpannya di dalam <i>database</i> |
| F19 | Penambangan Data | Modul dapat mendapatkan <i>frequent subgraph pattern</i> dari <i>social network</i> yang sudah tersimpan |

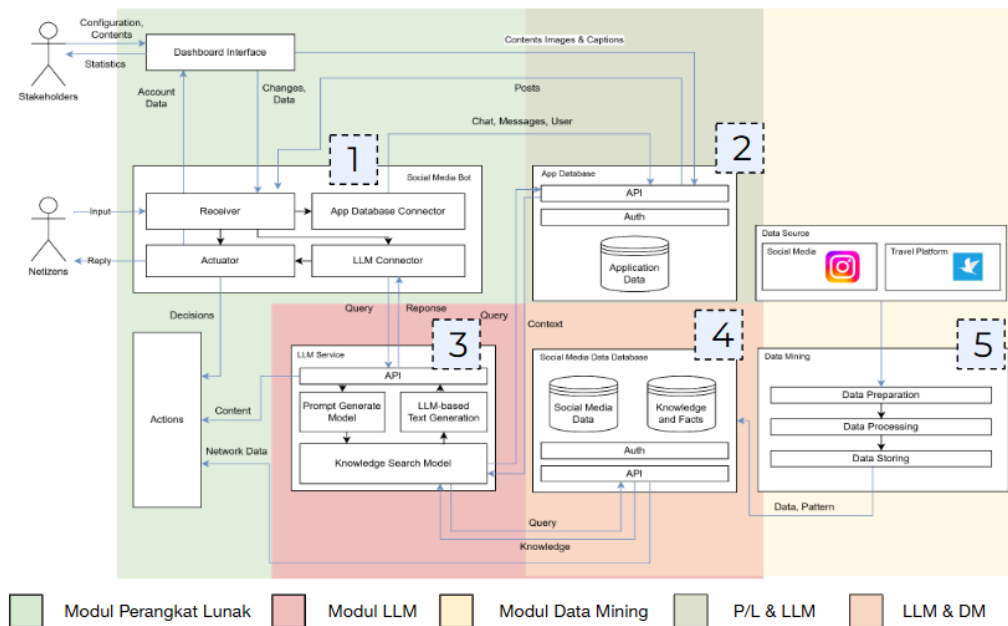
Berikut adalah spesifikasi dari kebutuhan non-fungsional sistem:

Tabel B.2 Spesifikasi kebutuhan non-fungsional sistem

| ID | Parameter | Kebutuhan |
|------|-----------------|--|
| NF01 | <i>Language</i> | Seluruh masukan maupun keluaran sistem menggunakan Bahasa Inggris. |

LAMPIRAN C. RANCANGAN SISTEM

Arsitektur sistem *social media bot* yang dikembangkan terdiri dari berbagai komponen yang terhubung satu sama lain. Komponen ini dapat merupakan bagian dari cakupan satu atau lebih modul. Terdapat lima komponen utama yang menyusun sistem *social media bot* sebagai sebuah kesatuan sistem. Berikut adalah diagram arsitektur dari *social media bot*:



Gambar C.1 Arsitektur diagram *social media bot*

Komponen-komponen dari arsitektur sistem dapat dijelaskan secara lebih rinci sebagai berikut:

Tabel C.1 Komponen utama *social media bot*

| No | Nama Komponen | Deskripsi Komponen | Cakupan Modul |
|----|-------------------------|--|----------------------|
| 1 | <i>Social Media Bot</i> | Komponen utama dalam sistem <i>social media bot</i> yang | Modul Sistem Otomasi |

| | | | |
|---|-----------------------------------|--|--|
| | | terhubung langsung dengan <i>platform</i> Instagram. | |
| 2 | <i>App Database</i> | Tempat penyimpanan catatan aksi dan interaksi yang dilakukan oleh <i>social media bot</i> . | Modul Sistem Otomasi, Modul <i>Dashboard</i> , dan Modul LLM |
| 3 | <i>LLM Service</i> | Komponen yang melakukan pemrosesan <i>input</i> menggunakan sistem LLMs dan pembelajaran mesin | Modul LLM |
| 4 | <i>Social Media Data Database</i> | Tempat penyimpanan pengetahuan dan pola yang didapatkan dari proses penambangan data. | Modul Penambangan Data dan Modul LLM |
| 5 | <i>Data Mining</i> | Komponen yang melakukan pemenuhan kebutuhan data untuk sistem <i>social media bot</i> | Modul Penambangan Data |

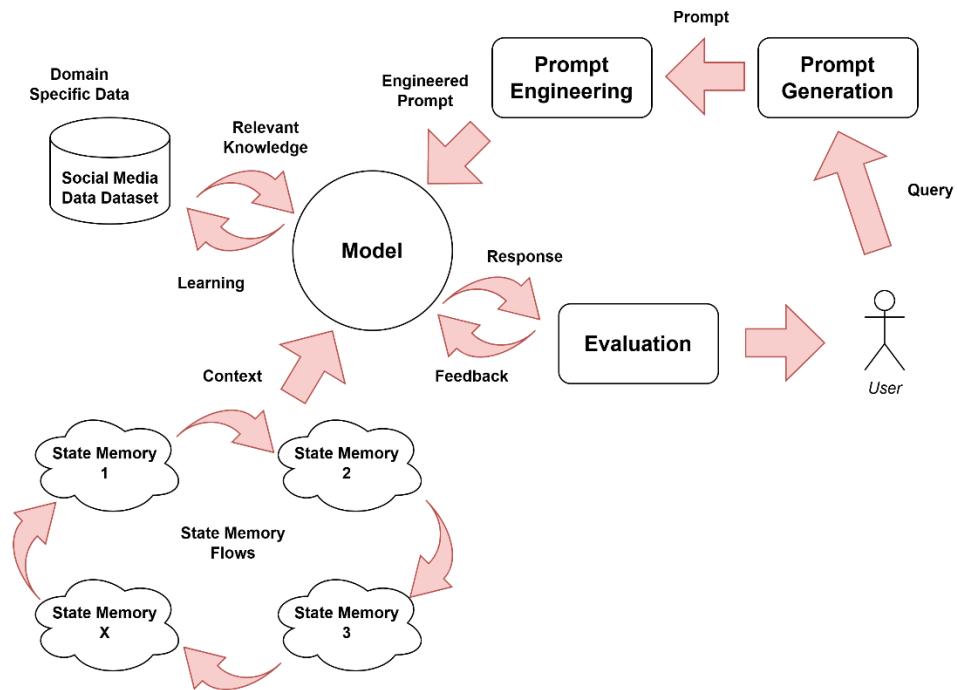
Diluar lima komponen utama, terdapat komponen-komponen penting lainnya dalam lingkungan pengembangan sistem *social media bot*, yakni:

Tabel C.2 Komponen pendukung *social media bot*

| Nama Komponen | Deskripsi Komponen | Cakupan Modul Utama | Cakupan Modul Pembantu |
|---------------|--------------------|---------------------|------------------------|
|---------------|--------------------|---------------------|------------------------|

| | | | |
|----------------------------|--|------------------------|--|
| <i>Dashboard Interface</i> | Antarmuka yang digunakan oleh Pihak <i>Stakeholders</i> untuk pelaksanaan konfigurasi dan pemenuhan data terhadap <i>bot</i> | Modul <i>Dashboard</i> | - |
| <i>Actions</i> | Aksi yang dikeluarkan oleh <i>bot</i> | Modul Sistem Otomasi | Modul LLM dalam pembangkitan konten untuk aksi <i>post</i> dan <i>comment</i> serta Modul Penambangan Data dalam penyediaan data jaringan dan koneksi untuk aksi <i>like</i> dan <i>follow</i> |
| <i>Data Source</i> | Sumber data relevan yang digunakan sebagai landasan pengetahuan <i>bot</i> | Modul Penambangan Data | - |

Bagian yang menjadi fokus dari pengerjaan tugas akhir ini adalah Modul LLM. Interaksi dan alur kerja antar komponen dalam Modul LLM dapat digambarkan dengan ilustrasi berikut:



Gambar C.2 Alur kerja sistem LLM *agent*

Penjelasan untuk setiap komponen dapat dijelaskan sebagai berikut:

Tabel C.3 Komponen sistem LLM *agent*

| Nama Komponen | Deskripsi | Interaksi Komponen |
|--------------------------|--|--|
| <i>Pre-trained Model</i> | Model yang menjadi basis bagi LLM yang akan diterapkan dalam <i>social media bot</i> | Berinteraksi dengan seluruh komponen lainnya sebagai basis dari pengembangan model final |
| Teknik <i>tuning</i> | Pelaksanaan peningkatan kinerja LLM dengan algoritma tertentu | Berinteraksi dengan Komponen <i>Pre-trained Model</i> untuk menghasilkan |

| | | |
|---------------------------|--|--|
| | | model final dan <i>database</i> untuk mendapatkan pengetahuan dan data yang relevan |
| Teknik <i>prompting</i> | Pelaksanaan pengarahannya instruksi kepada sistem LLMs | Berinteraksi dengan Komponen Teknik <i>Tuning</i> dalam pelaksanaan pembelajaran model, Komponen <i>State Memory Flows</i> dalam pemberian konteks yang benar, dan Komponen <i>Self-detection</i> untuk mengevaluasi <i>output</i> yang dihasilkan |
| <i>State memory flows</i> | Proses pergantian <i>state</i> dalam sistem LLMs dalam menyesuaikan konteks pembicaraan dengan pengguna | Berinteraksi dengan Komponen Teknik <i>Prompting</i> untuk memberikan konteks yang sesuai |
| <i>Self-detection</i> | Proses evaluasi terhadap <i>output</i> yang dikeluarkan oleh model agar sesuai dengan jawaban yang relevan | Berinteraksi dengan Komponen Teknik <i>Prompting</i> untuk pelaksanaan evaluasi atas <i>output</i> yang dikeluarkan model |